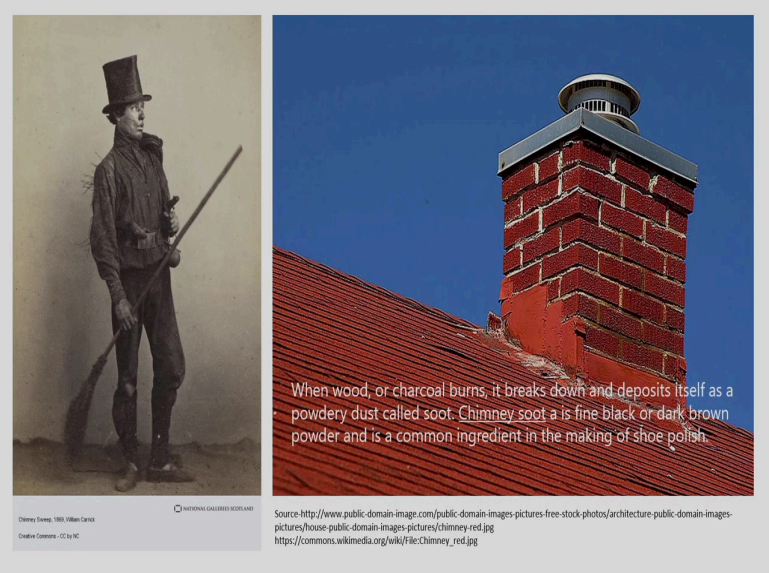
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 09
Genome Engineered Disease Modelling
Lecture - 01
Cancer Disease Models - Part A

Welcome to my course on Genome Editing and Engineering. Today we will be discussing module – 9, Genome engineered Disease Modelling, under which we will be discussing about Cancer disease models.

(Refer Slide Time: 00:51)

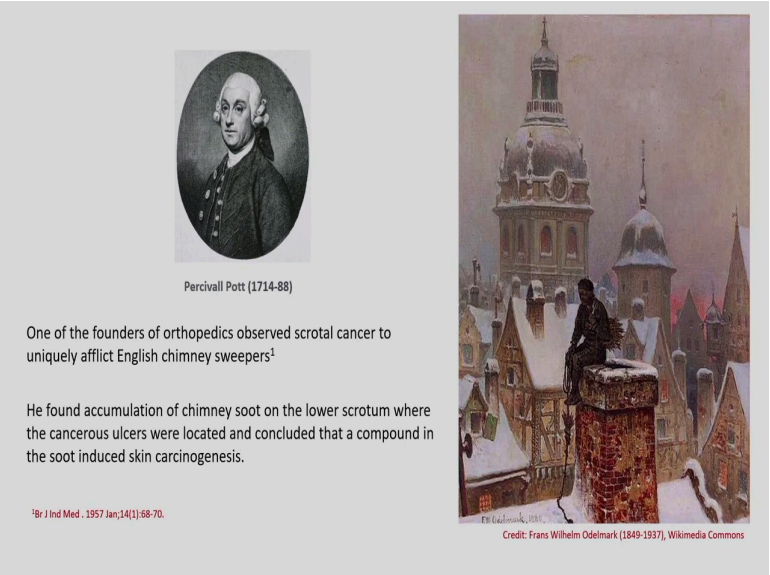


Before we start our lecture, let us have a look into this particular structure in a building and you can identify this is a chimney, which is atop a firewood place through which the smoke goes out of the building. And here wood and charcoal is burnt to produce heat and in that process it breaks down and deposits itself as a powdery dust, which we called as soot, it is black in color. And this chimney soot basically is a fine black and sometimes dark brown powder and this is a common ingredient in the making of shoe polish.

The chimneys get dirty due to the deposits of this soot and they have to be cleaned from time to time and you can see here a chimney cleaner, this photograph by William Carrick. And you

can see the darkened faces and hands and other body parts including the costume due to the chimney soot, which is stuck to the cleaner in the process while he was cleaning.

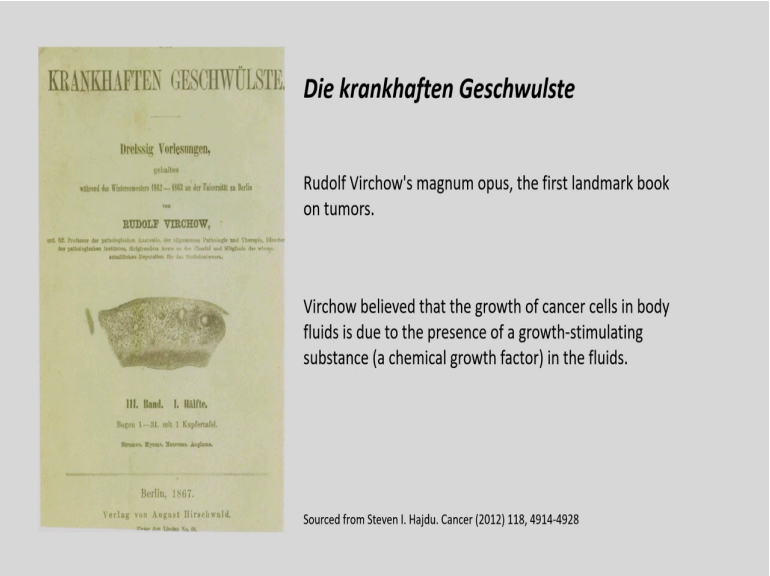
(Refer Slide Time: 02:02)



And you can see here another chimney cleaner sitting atop a chimney and this picture depicts a miserable person in a way. And here is one scientific study by Percivall Pott in 17, who is lived from 1714 to 88 and he is considered as the Founder of Orthopedics. And it is interesting that he observed a scrotal cancer to uniquely afflict the chimney sweepers in England.

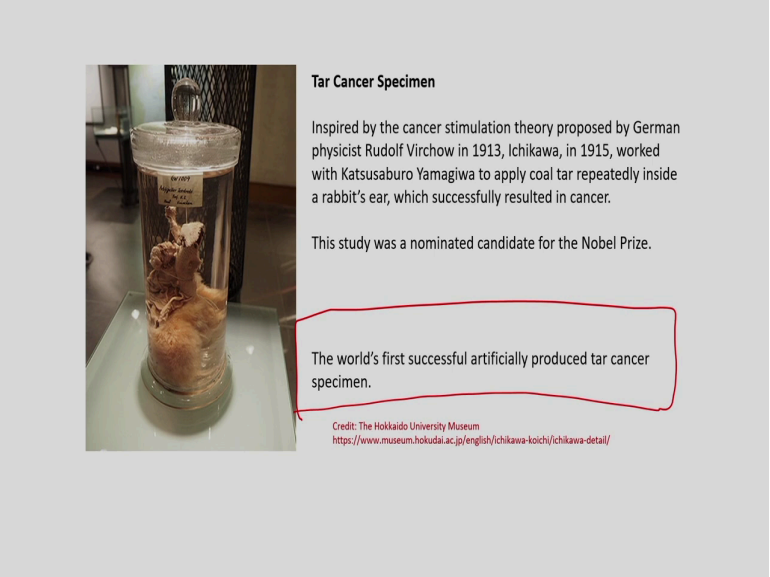
And he found the accumulation of chimney soot on the lower scrotum, where the cancerous ulcers were located and concluded that a compound in the soot probably induced skin carcinogenesis. This is an important fact that we need to discuss before you go move on to discuss about cancer modelling in cells and animals.

(Refer Slide Time: 03:05)



This is the magnum opus by Rudolf Virchow and the Krankhaften Geschwulste and this is the first landmark book on tumors. And Virchow believe that the growth of cancer cells in body fluids is due to the presence of a growth stimulating substance or a chemical growth factor, which exist in the fluids.

(Refer Slide Time: 03:31)

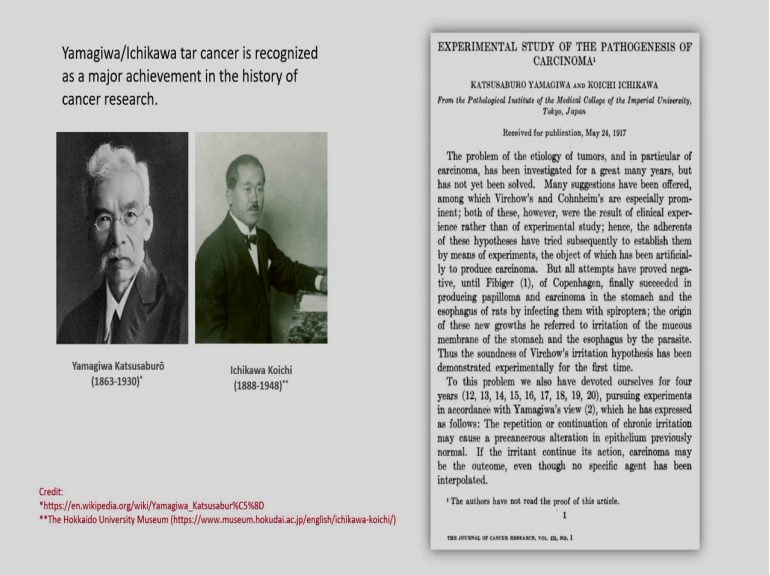


So, these are some of the important concepts and factual development regarding the understanding of cancer and you can see here a collection of a tar cancer specimen. And Virchow's study studies inspired Ichikawa in 1915, who along with Katsusaburo Yamigawa

applied coal tar repeatedly inside a rabbits ear, which successfully resulted in cancer and this study was one of the landmark studies which was nominated as a candidate for the Nobel Prize around that time.

This specimen is the world's first successful artificially produced tar cancer specimen. So, from the conceptual development of the chimney soot and the disease it causes in chimney cleaners to these artificially produced tar cancer specimen, we can see a common linkage that we can develop the disease model, which occurs in human subjects in animals.

(Refer Slide Time: 04:58)

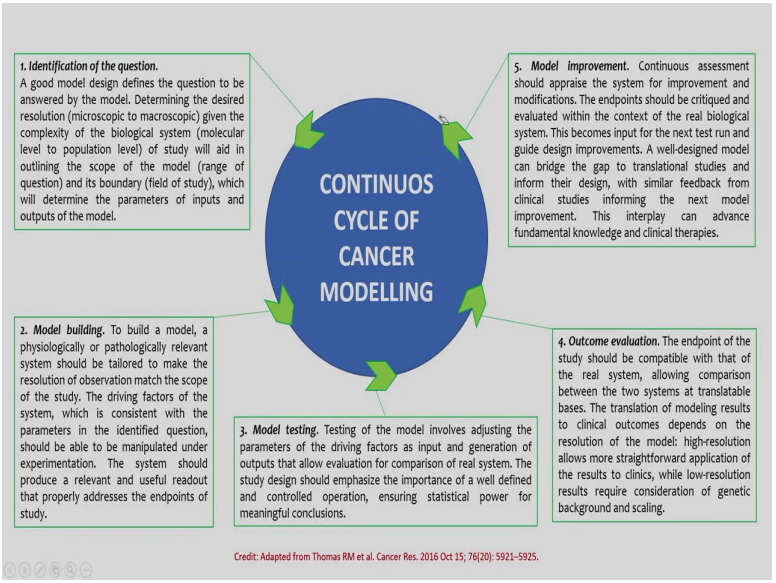


So, this study as I already told you, with tar inducing cancer in rabbits is recognized as a major achievement in the history of cancer research.

(Refer Slide Time: 05:10)

So, we have spoken to you about the disease modelling in the earlier lectures. Once again let us understand that modelling and experimental approach to investigate complex biological systems has significantly contributed to our understanding of cancer. Extensive cancer research has been conducted, utilizing animal models for elucidating mechanisms and developing therapeutics.

(Refer Slide Time: 05:42)



So, when we say a modelling of a disease and particularly cancer modelling, this is a continuous cyclic process. So, the first step in this entire process is identification of the

question. A good model design defines the question to be answered by the model, determines the desired resolution whether the studies to be done at the microscopic level or the macroscopic level.

Given the complexity of the biological system from molecular level to population level of study, will aid in outlining the scope of the model, which is the range of the question and its boundary the field of study, which will determine the parameters of input and outputs of the model.

The next step is model building; once we identify the question we go on to build up the model. To build a model a physiologically or pathologically relevant system should be tailored, to make the resolution of observation match the scope of the study. The driving factors of the system, which is consistent with the parameters in the identified question should be able to be manipulated under experimentation.

The system should produce a relevant and useful readout, that properly addresses the endpoints of the study. And in the third step we go for testing the model which we have built up in the second step. This testing of the model involves adjusting the parameters of the driving factors as input and generation of outputs that allows evaluation for comparison of real systems. The study design should emphasize the importance of a well defined and controlled operation, ensuring statistical power for meaningful conclusions.

Once the model testing is over, we go for evaluation of the outcome in the fourth step. The end point of the study should be compatible with that of the real system, allowing comparison between the two systems at translatable basis. The translation of modelling results to clinical outcomes depends on the resolution of the model; high resolution allows more straightforward applications of the results to clinics, while the low resolution results require consideration of genetic background and scaling.

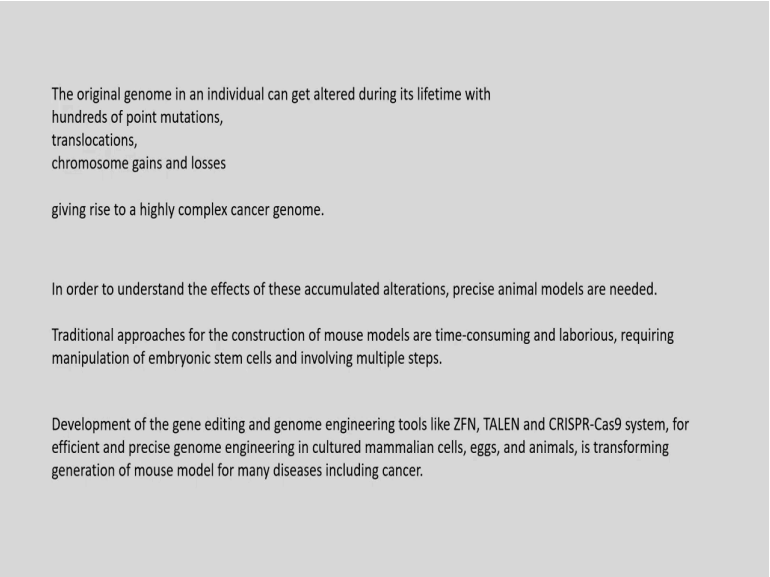
So, sometimes we need to improve the model. So, in the fifth step we go for model improvement, a continuous assessment should appraise the system for improvement and modifications. The endpoints should be critiqued and evaluated within the context of the real biological system, this becomes input for the next test run and guide the design improvement.

A well-designed model can bridge the gap to translational studies and inform their design with similar feedback from clinical studies informing the next model improvement. And this

can be a cyclic iterative process until and unless we get a model which is as good as the physical system.

This interplay can advance fundamental knowledge and clinical therapies. So, these are the main steps in the continuous cycle of cancer modelling: identification of the question, model building, model testing, outcome evolution and model improvement. And this is as you can see in the picture a cyclic process which can go on and on, until we get realistic results from the model.

(Refer Slide Time: 09:39)



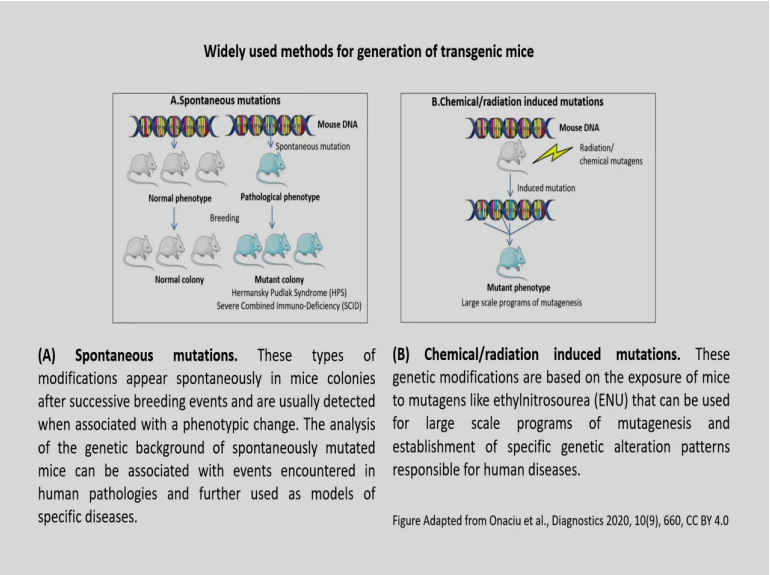
Let us now discuss about the genetics of cancer, what happens exactly at the genetic level or the to the genetic landscape whenever cancer occurs. The original genome in an individual can get altered during its lifetime, with hundreds of point mutations translocations, chromosome gains and losses as well as gene loss and gain.

And this gives rise to a highly complex cancer genome. In order to understand the effects of these accumulated alterations precise animal models are needed traditional approaches for the construction of mouse models are time consuming and laborious. And we have discussed about this animal modelling earlier and requires manipulation of embryonic stem cells and involves multiple steps.

Development of the gene editing and genome engineering tools like ZFN, TALEN and CRISPR-Cas9 system for efficient and precise genome engineering in cultured mammalian

cells, eggs, and animals, is transforming generation of mouse model and other animal models for many diseases including cancer.

(Refer Slide Time: 10:55)



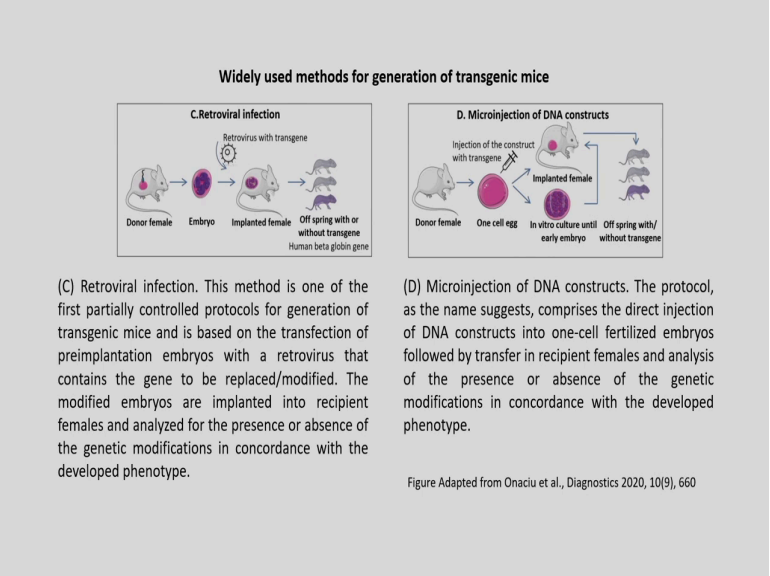
Let us briefly again discuss the widely used methods for generation of transgenic mice, with respect to model building. So, we may have spontaneous mutations or we may have chemical or radiation induced mutations. So, in these spontaneous mutations, which occurs due to natural mutations we may get a pathological phenotype from which we get a mutant colony.

And some of the animal models which can be generated or which are obtained through these are the Hermansky Pudlak syndrome or severe combined immuno deficiency or SCID. While in the second time, chemical or radiation induced mutations, we provide the radiation in the chemical mutations artificially, which induces some kind of genetic changes and gives rise to mutant phenotypes.

So, in the spontaneous mutations, modifications appear spontaneously in mice colonies, after successive breeding events and are usually detected when associated with a phenotypic change. The analysis of the genetic background of spontaneously mutated mice can be associated with the events encountered in human pathologies and further used as models of specific diseases.

While in chemical or radiation induced mutations, genetic mutations are based on the exposure of mice to mutagens that can be used for large scale programs of mutagenesis in establishment of specific genetic alteration patterns responsible for human diseases.

(Refer Slide Time: 12:46)



Other methods for generation of transgenic mice involves retroviral infection and micro injection of the DNA constructs. In micro viral infection partially controlled protocols for generation of transgenic mice is used and this is based on the transfection of pre implantation embryos, with a retrovirus that contains the gene to be replaced or modified.

The modified embryos are implanted into recipient females and analyzed for the presence or absence of the genetic modifications in concordance with the developed phenotype.

(Refer Slide Time: 13:48)

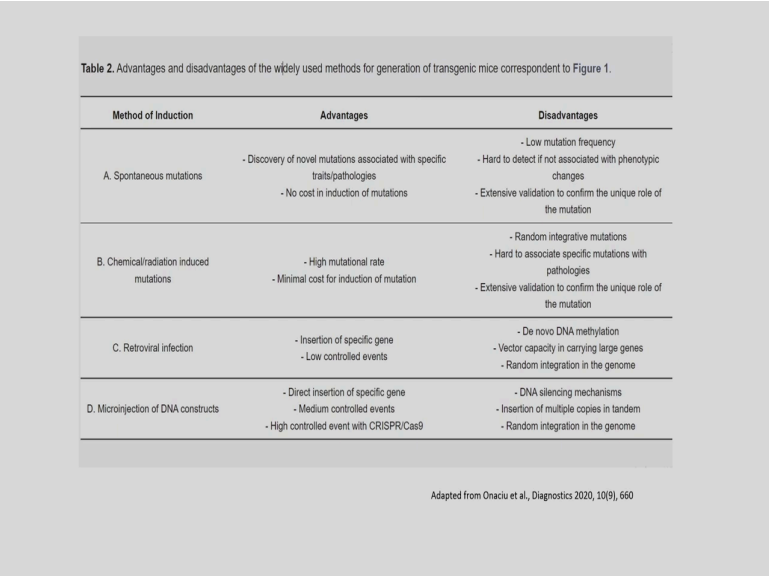


Table 2. Advantages and disadvantages of the widely used methods for generation of transgenic mice correspondent to Figure 1.

Method of Induction	Advantages	Disadvantages
A. Spontaneous mutations	- Discovery of novel mutations associated with specific traits/pathologies - No cost in induction of mutations	- Low mutation frequency - Hard to detect if not associated with phenotypic changes - Extensive validation to confirm the unique role of the mutation
B. Chemical/radiation induced mutations	- High mutational rate - Minimal cost for induction of mutation	- Random integrative mutations - Hard to associate specific mutations with pathologies - Extensive validation to confirm the unique role of the mutation
C. Retroviral infection	- Insertion of specific gene - Low controlled events	- De novo DNA methylation - Vector capacity in carrying large genes - Random integration in the genome
D. Microinjection of DNA constructs	- Direct insertion of specific gene - Medium controlled events - High controlled event with CRISPR/Cas9	- DNA silencing mechanisms - Insertion of multiple copies in tandem - Random integration in the genome

Adapted from Onaciu et al., Diagnostics 2020, 10(9), 660

While in the case of micro injection of DNA constructs, the protocol as the name suggest comprises the direct injection of DNA constructs into one cell fertilized embryos, followed by transfer in recipient females and analysis of the presence or absence of the genetic modifications in concordance with the developed phenotype. All these methods have certain advantages and disadvantages in the generation of transgenic mice.

For example, spontaneous mutations have the advantages like discovery of novel mutations associated with specific traits and pathologies and there is no cost in induction of the mutations. However, disadvantage is that such natural mutation frequencies are very low, hard to detect if not associated with phenotypic changes, extensive validation to confirm the unique role of the mutation is required.

In the case of chemical or in radiation induced mutations: we have high mutational rate, minimal cost for induction of the mutation, but disadvantages are that we have random integrative mutations, hard to associate specific mutations with pathologies, extensive validation to confirm the unique role of the mutation as in the case of spontaneous mutation.

In the case of retroviral infection, we have the advantage of insertion of specific genes and low control events. And disadvantage is the De novo DNA methylation, vector capacity in carrying large genes is limited and there is random integration in the genome. While in the case of micro injection of DNA constructs we have the advantage of direct insertion of specific genes, it is medium control events, high control event with CRISPR Cas9 utilization.

Disadvantage is the DNA silencing mechanism and the insertion of multiple copies in tandem, random integration in the genome.

(Refer Slide Time: 15:43)

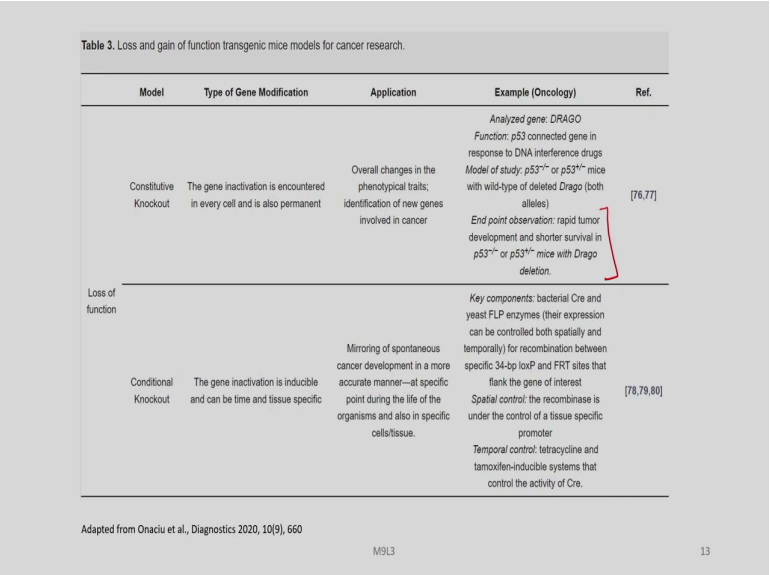


Table 3. Loss and gain of function transgenic mice models for cancer research.

	Model	Type of Gene Modification	Application	Example (Oncology)	Ref.
Loss of function	Constitutive Knockout	The gene inactivation is encountered in every cell and is also permanent	Overall changes in the phenotypical traits; identification of new genes involved in cancer	*Analyzed gene: DRAGO* *Function: p53* connected gene in response to DNA interference drugs *Model of study: $p53^{-/-}$* or *$p53^{+/-}$* mice with wild-type of deleted *Drago* (both alleles) *End point observation:* rapid tumor development and shorter survival in *$p53^{-/-}$* or *$p53^{+/-}$ mice with Drago deletion.*	[76,77]
	Conditional Knockout	The gene inactivation is inducible and can be time and tissue specific	Mirroring of spontaneous cancer development in a more accurate manner—at specific point during the life of the organisms and also in specific cells/tissue.	*Key components:* bacterial Cre and yeast FLP enzymes (their expression can be controlled both spatially and temporally) for recombination between specific 34-bp loxP and FRT sites that flank the gene of interest *Spatial control:* the recombinase is under the control of a tissue specific promoter *Temporal control:* tetracycline and tamoxifen-inducible systems that control the activity of Cre.	[78,79,80]

Adapted from Onaciu et al., Diagnostics 2020, 10(9), 660

M9L3 13

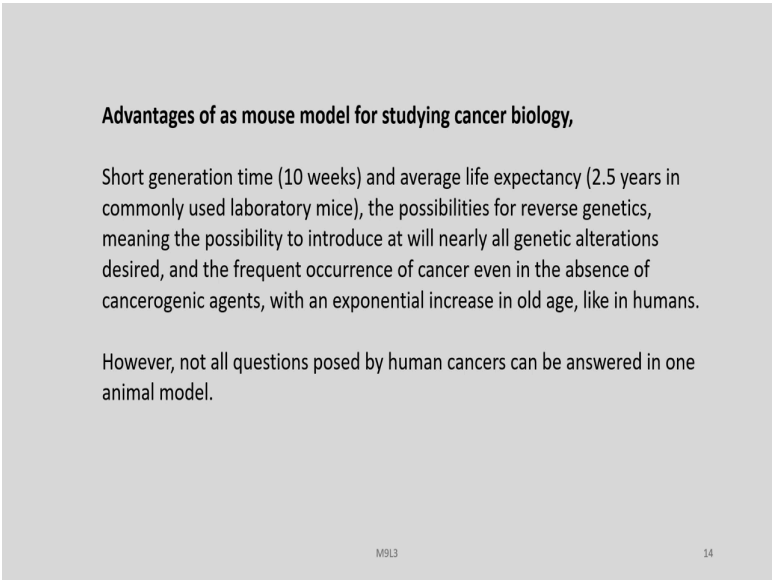
This table shows us a comparison between the loss and gain of function in transgenic mice models for cancer research. So, we may have models which may be for constitutive knockout models and then we may have conditional knockout models. In the case of constitutive knockout, the type of gene modification involves a gene activation which is encountered in every cell and is also permanent.

And this is applied in the case of overall changes in the phenotypical traits and identification of new genes involved in a cancer. And we have many examples like in the case of analyzing DRAGO function: p53 connected gene in response to DNA interference drugs and the model of the study p53 negative by negative or p53 positive by negative mice with wild type of deleted DRAGO in the both alleles.

And the end point observation in this case was rapid tumor development and shorter survival in p53 minus by minus or p53 plus minus mice with DRAGO deletion. In the case of conditional knockouts, the gene inactivation is inducible and can be time and tissue specific, mirroring of spontaneous cancer development in a more accurate manner. At specific point during the life of the organisms and also in specific cells and tissues.

Here, the key components are bacterial Cre and yeast FLP enzymes, their expressions can be controlled both spatially and temporarily for recombination between specific 34 base pair loxP and FRT sites that flank the gene of interest. Spatial control: the recombinase is under the control of a tissue specific promoter. And in temporal control tetra cycling and tamoxifen inducible systems that control the activity of Cre.

(Refer Slide Time: 18:12)

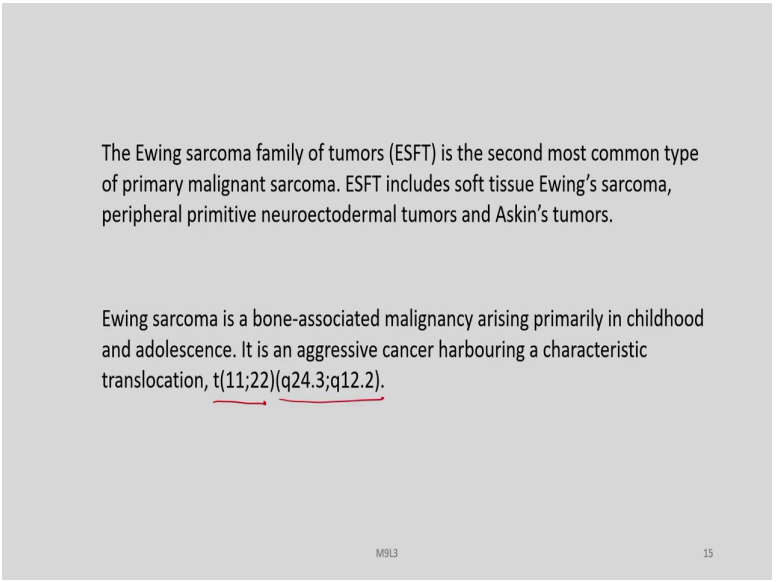


What are the advantages of mouse model for studying cancer biology? We have discussed about animal models in disease biology earlier, and we know that in the case of mouse the short generation time of 10 weeks and average life expectancy of 2.5 years in commonly used laboratory mice is one of the advantages.

And the possibilities of reverse genetics, meaning the possibility to introduce at will nearly all genetic alterations we desired and the frequent occurrence of cancer even in the absence of carcinogenic agents for example, due to spontaneous mutations, with an exponential increase in old age similar to those in humans are some of the additional advantages.

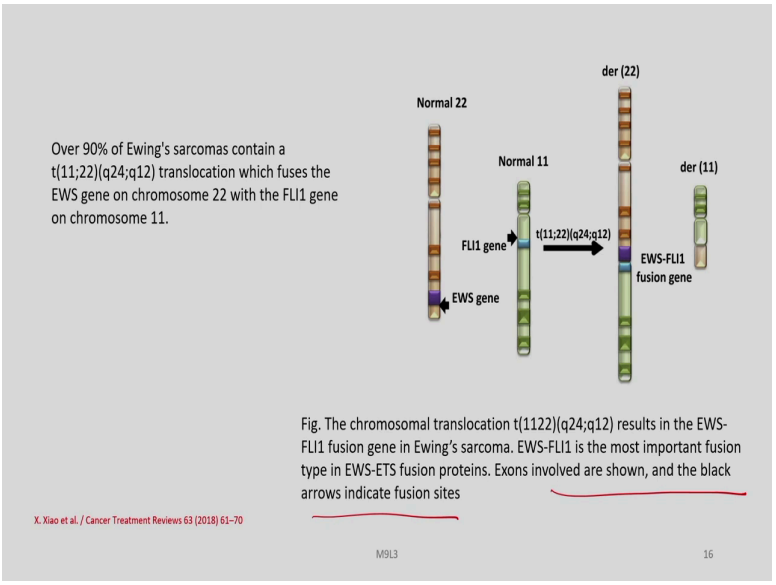
However, mouse is not an ideal animal in all cases, as not all questions posed by human cancers can be answered by mouse models alone and we may require models from other animals as well.

(Refer Slide Time: 19:12)



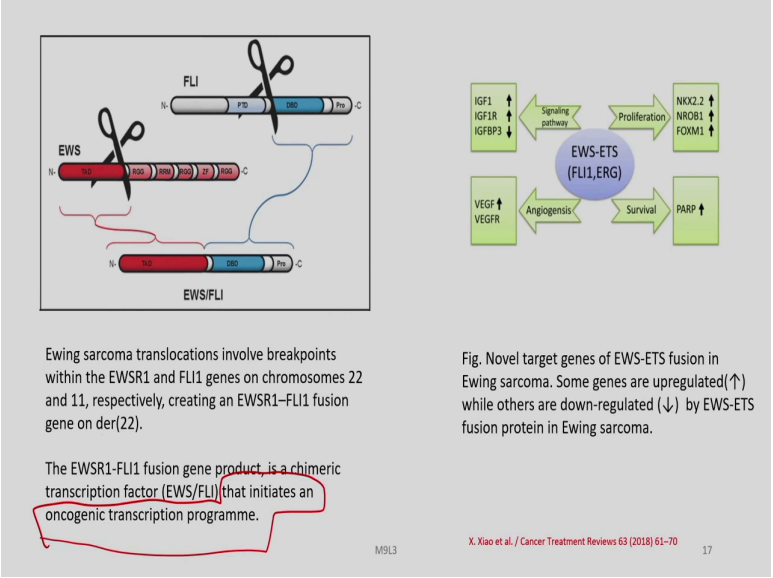
Let us now discuss about one specific case the Ewing sarcoma family of tumors (ESFT), which is the second most common type of primary malignant sarcoma. Ewing sarcoma family of tumors includes soft tissue Ewing sarcoma, peripheral primitive neuroectodermal tumors and Askin's tumors. Ewing sarcoma is a bone associated malignancy, which arise primarily in childhood and adolescence. It is an aggressive cancer harbouring a characteristic translocation t(11 is to 22)(q24.3 is to q12.2).

(Refer Slide Time: 20:05)



Over 90 percent of Ewing sarcomas contain a t(1122)(q24;q12) translocation which fuses the EWS gene on chromosome 22 with FLI1 gene on chromosome 11. So, this is the EWS-FLI1 fusion gene here is you can see here. The chromosomal translocation here results in the EWS-FLI fusion gene as already shown to you in the figure. This EWS-FLI1 is the most important fusion type in EWS-ETS fusion proteins. Exons involved are shown, and the black arrows indicate the fusion sites in this picture.

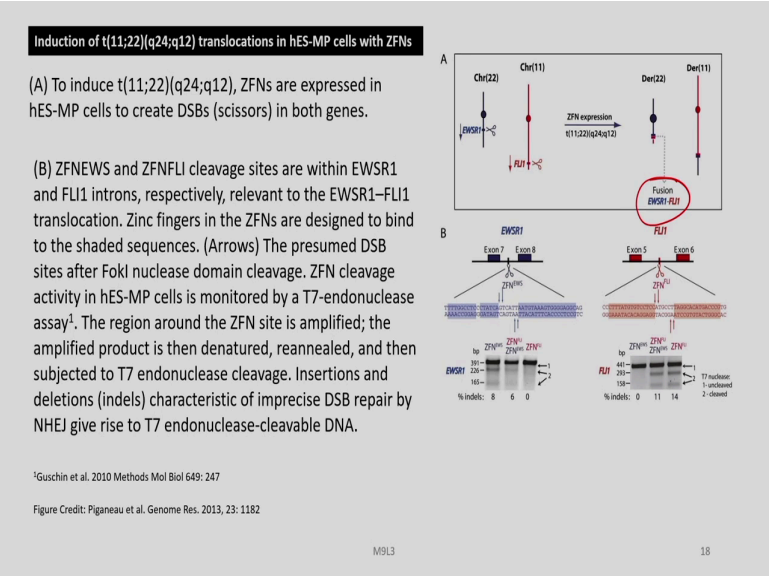
(Refer Slide Time: 21:07)



So, the Ewing sarcoma translocations involves breakpoints within the EWSR1 and FLI1 genes on chromosome 22 and 11, respectively, which creates an EWSR1 and FLI1 fusion gene on der(22). The EWSR1-FLI1 fusion gene product, is a chimeric transcription factor, EWS FLI that initiates an oncogenic transcription programme.

So, this is the EWS-ETS FLI1, ERG. So, you can see impacting signaling pathways proliferation, then angiogenesis as well as survival. Certain genes are up regulated as shown by the upward arrow and certain other genes are down regulated as shown by the downward arrow. So, the novel target genes of EWS ETS fusion in Ewing sarcoma you can see here in this figure, involved in proliferation survival angiogenesis and signaling pathway. So, this translocation can be artificially inducted with the help of various techniques.

(Refer Slide Time: 22:48)

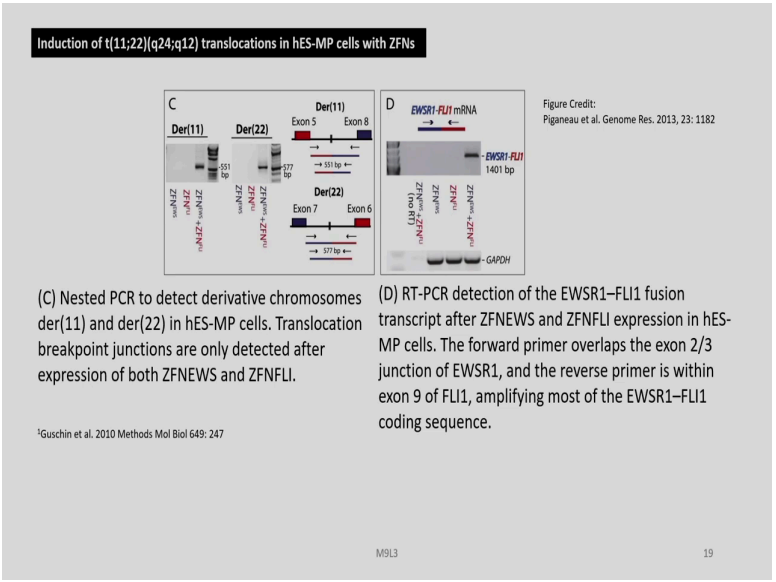


So, here we are inducing this translocation in hES-MP cells, with the help of zinc finger nucleases. So, you can see here the two chromosomes 22 and chromosome 11 and then there is the application of ZFN, which produces the fusion EWSR1 and FLI1 as shown in this original schematic diagram.

So, to induce these fusion genes the ZFNs are expressed in hES-MP cells to create double strength breaks in both the genes as you can see here. The ZFNEWS and ZFNFLI cleavage sites are within the EWSR1 and FLI1 introns respectively, relevant to the EWSR1-FLI1 translocation.

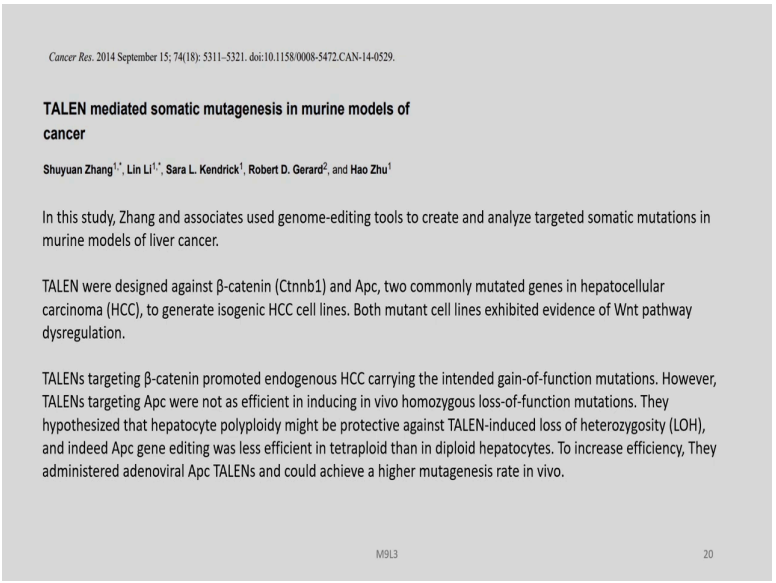
Since zinc fingers in the ZFNs are designed to bind to the shaded sequence the arrows are shown there the presumed DSB sites after Fok1 nucleus domain cleavage. ZFN cleavage activity in hES-MP cells is monitored by a T7 endonuclease assay. The region around the ZFN site is amplified the amplified product is then denatured reannealed and then subjected to T7 endonuclease cleavage, insertions and deletions indels characteristic of imprecise DSB repair by NHEJ give rise to T7 endonuclease cleavable DNA.

(Refer Slide Time: 24:47)



Nested PCR was used to detect derivative chromosomes, the der11 and der22 in hES-MP cells translocation breakpoint junctions are only detected after expression of both ZFN-EWS and ZFN-FLI. RT-PCR detection of the EWSR1 and FLI1 fusion transcript can be see in figure D, after the ZFN-EWS and ZFN-FLI expression in hES-MP cells. The forward primer overlaps the exon 2 by 3 junction of EWSR1 and the reverse primer is within exon 9 of FLI1 amplifying most of the EWSR1-FLI1 coding sequence.

(Refer Slide Time: 25:32)

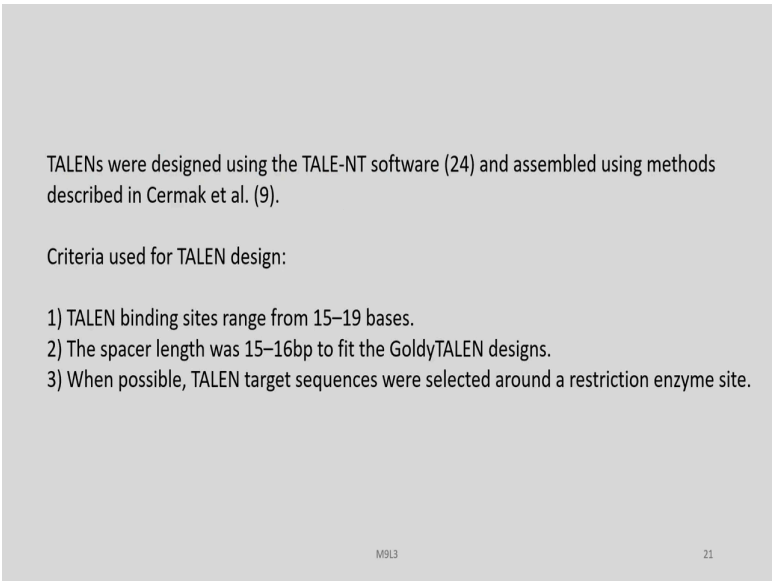


Apart from using ZFN to create such disease models, TALEN has been also used successfully for somatic mutagenesis in creating murine models of cancer. In this work by Zhang et al, they used genome editing tool TALEN to create and analyze targeted somatic mutations in a murine models of a liver cancer. The TALENs were designed against beta catenin and Apc, two commonly mutated genes in hepatocellular carcinoma or HCC to generate isogenic HC cell lines. Both mutant cell lines exhibited evidence of Wnt pathway dysregulation.

TALENs targeting beta catenin promoted endogenous HCC carrying the intended gain of function mutations. However, TALENs targeting Apc were not as efficient in inducing in vivo homozygous loss of function mutations. They hypothesize that hepatocyte polyploidy might be protective against TALEN induced loss of heterozygosity and indeed Apc gene editing was less efficient in tetraploid than in diploid hepatocytes.

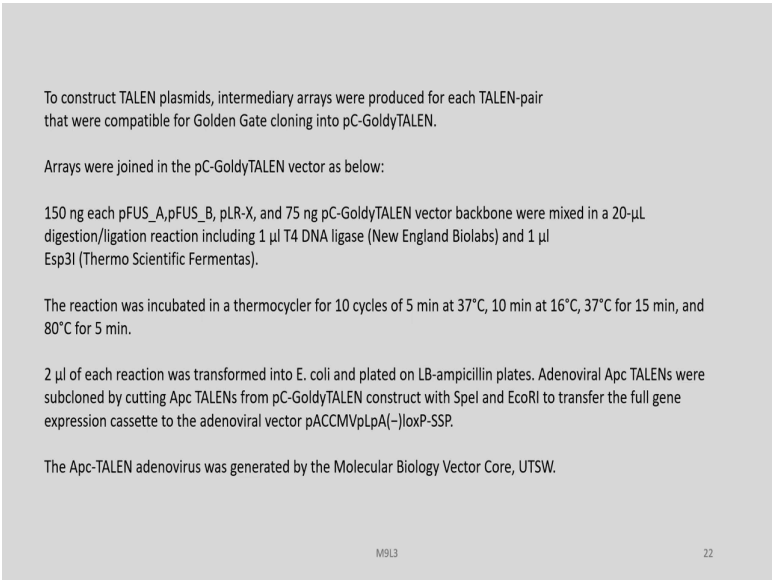
To increase efficiency, they administered adenoviral Apc TALENs and could achieve a higher mutagenesis rate in vivo.

(Refer Slide Time: 27:01)



A TALENs were designed using the TALE-NT software, the criteria used for TALEN design are as bellows: TALEN binding sites range from 15 to 19 bases. The spacer length was 15 to 16 base pairs to fit the GoldyTALEN designs. When possible, TALEN target sequences were selected around a restriction enzyme site.

(Refer Slide Time: 27:26)

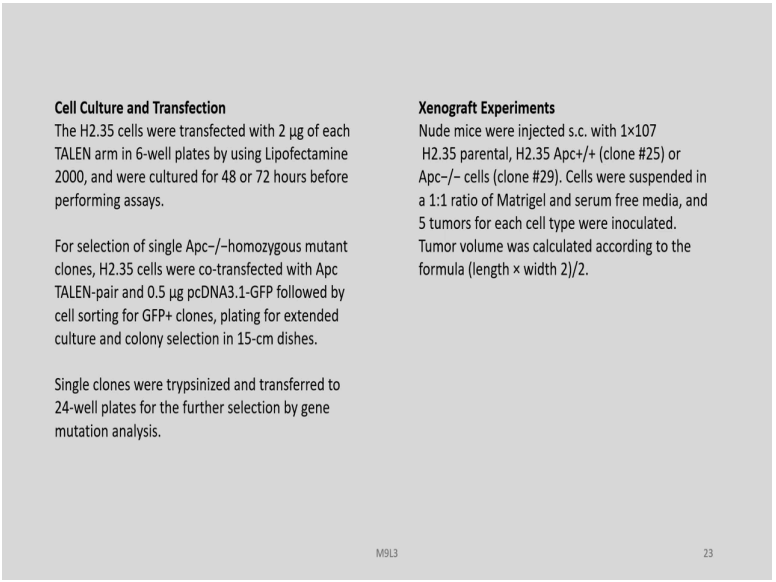


To construct TALEN plasmids, intermediary arrays were produced for each TALEN-pair that were compatible for Golden Gate cloning into pC-GoldyTALEN. Arrays were joined in the pC-GoldyTALEN vector as below. 150 nano gram each pFUS A, pFUS B, pLR-X and 75 nanogram pC-GoldyTALEN vector backbone were mixed in the 20 micro liter digestion ligation reaction, including 1 micro liter T4 DNA ligase and 1 micro liter Esp31.

The reaction was incubated in thermocycle for 10 cycles of 5 minute at 37 degree centigrade 10 minute at 16 degree centigrade, 37 degree for 15 minute and 80 degree centigrade for 5 minute. 2 micro liter of each reaction was transformed into E coli and plated on LB-ampicillin plates.

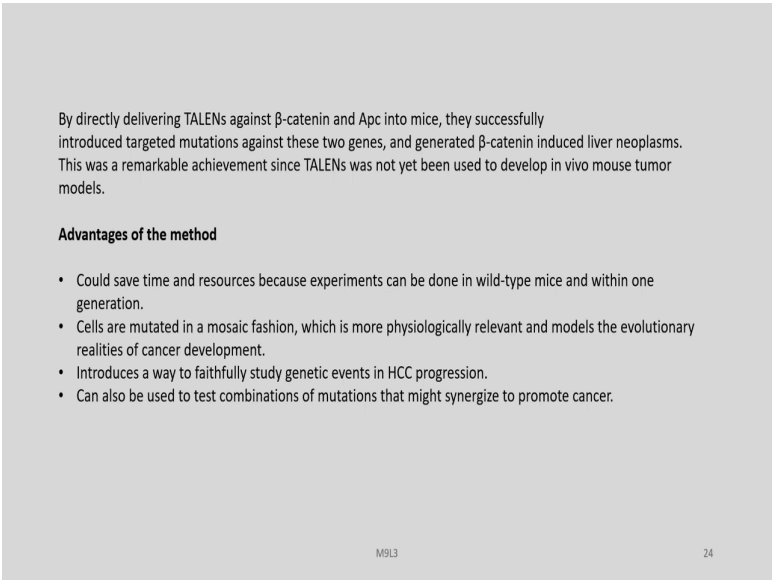
Adenoviral Apc TALENs were subcloned by cutting Apc TALENs from pC-GoldyTALEN construct with Spe 1 and EcoR1 to transfer the full gene expression cassette to the adenoviral vector pACCMVpLpA(-)loxP-SSP. The Apc-TALEN adinovirus was generated by the Molecular Biology Vector Core, UTSW.

(Refer Slide Time: 28:50)



This was followed by cell culture and transfection and finally, xenograft experimentation. The nude mice were injected with 1 into 107 H2.35 parental, H2.35 Apc+/+ clone or Apc-/- cells. Cells were suspended in a 1 is to 1 ratio of Matrigel and serum free media and 5 tumors for each cell type were inoculated. Tumor volume was calculated according to the formula length into width 2 divided by 2.

(Refer Slide Time: 29:30)



By directly delivering TALENs against beta catenin and Apc into mice, Zhang et al successfully introduced Zhang et al successfully introduced targeted mutations against these

two genes, and generated beta catenin induced liver neoplasms. This was a remarkable achievement since TALENs was not yet used at the time to develop in vivo mouse tumor models.

Some of the advantage of the method developed by Zhang and associates are it could save time and resources because experiments can be done in wild type mice and within one generation. Cells are mutated in a mosaic fashion, which is a more physiologically relevant and models the evolutionary realities of cancer development.

It introduces a way to faithfully study genetic events in HCC progression. It can also be used to test combinations of mutations that might synergize to promote a cancer. With this we come to the basics of a cancer models and the importance of cancer models and also two of the gene editing technologies like ZFN and TALEN in producing cancer models in as in cells; so.

Thank you for your patient hearing. We will be continuing this lecture in Part B.